Simulated Annealing with Tsallis Weights
A Numerical Comparison

Ulrich H.E. Hansmann[1]
Department of Theoretical Studies
Institute for Molecular Science
Okazaki, Aichi 444, Japan

ABSTRACT

We discuss the use of Tsallis generalized mechanics in simulated annealing algorithms. For a small peptide it is shown that older implementations are not more effective than regular simulated annealing in finding ground state configurations. We propose a new implementation which leads to an improvement over regular simulated annealing.

Keywords: Simulated Annealing, Tsallis Statistics, Protein Folding

[1] e-mail: hansmann@ims.ac.jp

Simulated annealing [1] has become an often used tool to tackle hard optimization problems in various fields of science. Its underlying idea of modeling the crystal grow process in nature is easy to understand and simple to implement. Any Monte Carlo or molecular dynamics technique can be converted into a simulated annealing algorithm by allowing the temperature to decrease gradually. However, the method is not without problems. The performance of simulated annealing depends crucially on the annealing schedule. It could be shown that convergence to the global minimum can be secured for a logarithmic annealing schedule [2], but this is of little use in applications of the method. Constraints in available computer time enforce the choice of faster annealing schedules where success is no longer guaranteed. Similar to the growing of real crystals which hardly ever can be archived by a simple cooling process, elaborated and often system specific annealing schedules are frequently necessary to obtain the global minimum in the CPU time available. Hence, ever since the seminal article by Kirkpatrick *et al.* [1], attempts were made, to improve the performance of simulated annealing in practical applications, see for instance Ref. [3]. Two of the more recent attempts [4, 5] are inspired by Tsallis generalized mechanics [6, 7]. We compare the performance of both implementations with regular simulated annealing by taking an energy function for the protein folding problem as an example. We describe then a new application of Tsallis weights to simulated annealing which leads for our model to an improvement over regular simulated annealing.

In the Tsallis formalism [6], a generalized mechanics is constructed by maximizing a generalized entropy

$$S = -k \sum_i p_i \ln p_i \tag{1}$$

with the constraints

$$\sum_i p_i = 1 \qquad \sum_i p_i^q E_i = \text{const} \ . \tag{2}$$

Here, q is a real number. A generalized probability distribution

$$p_i \propto [1 - (1-q)\beta E_i]^{\frac{1}{1-q}} \tag{3}$$

follows and the average of an observable $\mathcal{O}$ can be defined by

$$< \mathcal{O} >= \sum_i p_i^q \mathcal{O}_i \ . \tag{4}$$

It is evident that for $q \to 1$ the generalized distribution of Eq. 3 tends toward the Gibbs-Boltzmann distribution and therefore regular statistical mechanics is recovered in this limit.

The important feature of Tsallis generalized statistic for optimization problems is that the probability of states does no longer decrease exponentially with energy but according to a power law where the exponent is determined by the free parameter q (see Eq. 3). This observation inspired a generalized simulated annealing algorithm [4] where the acceptance probability

$$p(\Delta E) = \min\left[1, [1-(1-q)\beta\Delta E]^{\frac{1}{1-q}}\right] \tag{5}$$

was introduced. Here, ΔE is the change in energy. Again, for $q \to 1$, the acceptance probability of canonical simulated annealing is recovered. The algorithm was employed to find close to optimal solutions to the traveling salesman problem and it was claimed that it is faster than classical simulated annealing and has optimal performance for large negative parameters q [4]. However, the above algorithm does not obey detailed balance and therefore convergence to an equilibrium distribution is in general not guaranteed. For this reason, it was recently proposed to utilizes the acceptance probability [5]

$$p(E_{old} \to E_{new}) = \min\left[1, \left(\frac{1-[1-q(T)]\beta E_{new}}{1-[1-q(T)]\beta E_{old}}\right)^{\frac{q(T)}{1-q(T)}}\right] \tag{6}$$

with $\lim_{T\to 0} q(T) = 1$. Detailed balance is obtained by the algorithm and convergence to the generalized distribution of Eq. 3 is guaranteed for each temperature T and parameter q. Since $q \to 1$ as $T \to 0$, this generalized simulated annealing algorithm tends in the same way as regular simulated annealing to a steepest descent at low temperature. For Tsallis parameters $q(T) > 1$ the distribution of Eq. 3 has a tail to higher energies and the probability to cross energy barriers and to escape local minima is therefore increased by the above weights. The new algorithm was applied in Ref. [5] to the conformational optimization of the 48 atom tetraalanine peptide using molecular dynamics and hybrid MD-MC methods in the CHARMM force field [8]. Both temperature and the Tsallis parameter q were exponentially decreased, with a start value of $q = 2$ for the Tsallis parameter. Results better than conventional simulated annealing were reported.

However, a disadvantage of the above algorithm is that it requires the careful tuning of additional free parameters. Not only a suitable annealing schedule in temperature T has to be chosen, but also one in the Tsallis parameter $q(T)$. Furthermore, the performance of this algorithm depends also on the choice of the zero in potential energy (which was ignored by the authors of Ref. [5]). This is because the transition probabilities of Eq. 6 are not invariant under a shift in energy. For instance, even for $q > 1$ (the case $q < 1$ can lead to complex probabilities) the weights of Eq. 6 can become negative for negative values of the energies. On the other hand, acceptance of a configuration will no longer depend on the energy of the configuration if all energies are shifted by a large enough positive number. Hence, while classical simulated annealing has already the problem of finding an optimal annealing schedule for the temperature, the above algorithm requires in addition a careful tuning of the Tsallis parameter q and of the energy scale, making its applicability highly model dependent.

Here we show how the above problems can be alleviated. Use of Tsallis weights in the course of a simulated annealing simulation is motivated by the fact that the resulting probability distribution has a tail to higher energies for $q > 1$, enhancing in this way the probability to cross barriers and escape local minima. While negligible at high temperatures this feature becomes important at low temperatures where a canonical weight would make it difficult to escape local minimas. We are therefore mainly interested in the use of of Tsallis weights for *low* temperatures. It is obvious that the Tsallis distribution at low temperatures should not be dominated by the tail to higher energies, but still be centered around the energy where the canonical distribution has its maximum. For otherwise, low energy (temperature) states will not be sampled sufficiently. To find the parameter $q > 1$ which yields to an optimal distribution let us first write the Tsallis weights as

$$w(E) = [1 - (1 - q)\beta(E - E_0)]^{\frac{q}{1-q}} \ , \tag{7}$$

where E_0 is the (in general unknown) ground state energy. This is equivalent to chosing an energy scale where all energies are positive with the zero for the ground state. In this way we ensure that the weights are always positive. The Tsallis weights will be a good approximation of the Boltzmann weights $W_B(E) = \exp(-\beta(E - E_0))$ for

$(1-q)\beta(E-E_0) << 1$. To ensure that simulations are able to escape from energy local minima, the weights should start deviating from the exponentially damped Boltzmann weights at energies near its mean value. This is because at low temperatures there are only small fluctuations of energy around its mean. We may thus set in Eq. 7:

$$-(1-q)\beta\left(< E > \; - \; E_0\right) = \frac{1}{2} \tag{8}$$

The mean value of energy is given at low temperatures by the harmonic approximation:

$$< E > -E_0 \approx \frac{n_F}{2} kT = \frac{n_F}{2\beta} \; , \tag{9}$$

where n_F is the degree of freedom of our molecule. Hence, for low temperatures, Eq. 8 can be written, as

$$-(1-q)\frac{n_F}{2} = \frac{1}{2} \; , \tag{10}$$

which leads to an optimal Tsallis parameter

$$q = 1 + \frac{1}{n_F} \; . \tag{11}$$

Hence, we propose a generalized simulated annealing algorithm where configurations are weighted with

$$w(E) = \left[1 - (1-q)\,\beta\,(E-E_0)\right]^{\frac{q}{1-q}} \; . \tag{12}$$

The Tsallis parameter q is set to $q = 1 + 1/n_F$. Our weights require knowledge of the ground state energy. However, in general E_0 is not known. We therefore approximate E_0 in the course of a simulated annealing simulation by $E_0 \equiv E_{min} - c$ where E_{min} is the lowest energy ever encountered in the simulation and c a small number. E_0 is reset every time a new value for E_{min} is found. Changing the value of E_0 is a disturbance of the Markov chain and while we expect the disturbance to be small, we clearly cannot use our algorithm to calculate thermodynamic averages. However, due to finite stepsize of the temperature annealing we can anyway not assume convergence against an equilibrium distribution. As with regular simulated annealing, our method is valid only as a global optimization method.

We have tested the various simulated annealing algorithms for the protein folding problem, a long-standing problem in biophysics with rough energy landscape. Here, Met-enkephalin has become a often used model to examine new algorithms. Met-enkephalin

has the amino acid sequence Tyr-Gly-Gly-Phe-Met. The potential energy function E_{tot} that we used is given by the sum of electrostatic term E_C, Lennard-Jones term E_{LJ}, and hydrogen-bond term E_{hb} for all pairs of atoms in the peptide together with the torsion term E_{tors} for all torsion angles. The parameters for the energy function were adopted from ECEPP/2.[9] Fixing the peptide bond angles ω to 180° leaves us with 19 torsion angles as degree of freedom. The computer code KONF90 [10] was used.

As in earlier work on Met-enkephalin [11] we performed for each of the different algorithms 20 runs of 50,000 sweeps. Each run started from completely random configuration and each angle is updated once in a sweep. The temperature was lowered exponentially according to

$$T = T_{ST}\gamma^{i-1} \tag{13}$$

for the ith sweep and

$$\gamma = (T_{FI}/T_{ST})^{\frac{1}{49999}} \ , \tag{14}$$

where T_{ST} is the start temperature and T_{FI} is the final temperature. One of the quantities we monitored to evaluate the performance of the various algorithms was the average $< E_{Low} >$ (taken over all 20 runs) of the lowest energies E_{Low} obtained in each single run. The other quantity was the number n_G of ground-state configurations found in the 20 independent runs. In Ref. [12] it was shown that with the energy function KONF90, conformations of energy less than -11.0 kcal/mol have essentially the same three-dimensional structure. Hence, we consider any conformation with $E \leq -11.0$ kcal/mol as the ground-state configuration.

Let us present now our results. In Tab. 1 we show typical results for the first generalized simulated annealing algorithm which was proposed in Ref. [4] and uses the acceptance probability of Eq. 5. We chose as start temperature $T_{ST} = 1000\ K$ and tried two values for the final temperature: $T_F = 50\ K$ and $T_F = 1\ K$. These temperatures were also chosen by us in Ref. [11] from which we have also taken the results for regular simulated annealing (indicated by $q = 1$ in Tab. 1). We did not observe an improvement over regular simulated annealing for any choice of the Tsallis parameter q in this generalized simulated annealing algorithm. In the range $0 \leq q \leq 1.25$ the performance is comparable with canonical simulated annealing. Both n_G and $< E_{Low} >$ vary little in this range. The

performance of the algorithm detoriates quickly outside of this range. This is especially true for large negative values of q which were presented as the optimal choice for q in Ref. [4]. The probability of finding ground state configurations becomes small and the average $< E_{Low} >$ of lowest energies found in each single run not only increases, but also the standard deviation of $< E_{Low} >$ indicating that the obtained low energies strongly depend on the initial configuration and the algorithm reduce in this case to a mere quenching. Hence, for our system, application of this generalized simulated annealing algorithm seems to bring no improvement over regular simulated annealing.

In Tab. 2 we show our results for the second method, which uses the acceptance probability of Eq. 6 and was proposed in Ref. [5]. Following the authors of Ref. [5] we chose the start value $q = 2$ and decreased both temperature and Tsallis parameter exponentially such that for the final temperature T_{FI} we have $q(T_{FI}) = 1$. Again we tried two values for the final temperature: $T_F = 50\ K$ and $T_F = 1\ K$. Since Tsallis distributions have a tail to high energies for $q > 1$, we expected that it would be possible to choose lower start temperatures than for regular simulated annealing. Hence, we tried not only $T_{ST} = 1000\ K$, but also $T_{ST} = 500\ K$ and $300\ K$. In each case we found a poorer performance than for regular simulated annealing. This is in contradiction to the results of Ref. [5], where the authors found a significant improvement over canonical simulated annealing. We remark that for the calculation of weights (see Eq. 6) in the simulations, we shifted the energies by the ground-state energy for Met-enkephalin (as known from previous work [11]) to ensure positive weights, otherwise even poorer results were obtained. Hence, the poor performance of this algorithm in simulations of our peptide is not only due to a poor choice of the energy scale, but also to an imperfect annealing schedule of the Tsallis parameter $q(T)$. We conclude that the performance of this algorithm is highly model dependent and requires carefully tuning in the annealing of q and the choice of the energy scale. This is a severe limitation for practical applications.

Finally, Tab. 3 shows the results for our implementation of Tsallis weights in simulated annealing algorithms using the acceptance probability of Eq. 12. The main difference to the previous algorithm is that the Tsallis parameter q is not free, but set to an optimal value of $q_F = 1 + 1/n_F = 1 + 1/19 \approx 1.053$. In addition, our procedure guarantees that

the weights are always positive by self-tuning the estimate of the ground state energy E_0 in the course of an annealing run. E_0 is reset every time to $E_0 = E_{min} - 1\ kcal/mol$ when a new configuration with lower energy E_{min} than any previous configuration is found. We found for both canonical and generalized simulated annealing an optimal performance for start temperature $T_{ST} = 500\ K$ and final temperature $T_F = 50\ K$. With this temperature annealing schedule the ground state configuration was found 8 out of 20 runs for regular simulated annealing and 12 out of 20 runs for generalized simulated annealing. This is a modest improvement of the new algorithm over the canonical simulated annealing. The improvement can also be seen in the estimate for $< E_{Low} >$ which is 0.6 kcal/mol lower for the new algorithm and has a smaller standard deviation than regular simulated annealing. We further notice that the new generalized ensemble algorithm allows to start the temperature annealing at lower temperatures. While regular simulated annealing works best with start temperatures over 500 K, the performance of the new algorithm depends only little on the start temperature and rather favors $T_{ST} \leq 500\ K$. This follows from the form of the Tsallis distributions which have a tail to high energies for $q > 1$. Equilibrization at lower temperatures is therefore enhanced. We remark that the improvement of the new method over regular simulated annealing still does not lead to the performance reported for the generalized ensembles algorithms in Refs. [11, 13]. However, the new algorithm is much easier to implement. Since the Tsallis parameter q is constant in our algorithm, only minor modifications are required in existing simulated annealing programs to accommodate the new technique. Unlike in the algorithm of Ref. [5] the improvement over regular simulated annealing is gained without the need of determining optimal annealing schedules for additional parameters.

Let us summarize our results. We have performed Monte Carlo simulations of Met-enkephalin using Tsallis generalized Mechanics. We discussed older proposals for the use of Tsallis weights in simulated annealing algorithms and showed how to overcome their shortcomings. Our algorithm is easy to implement in existing simulated annealing programs and does not require tuning of annealing schedules in additional variables. For the case of a simple peptide we have demonstrated that our new technique offers an easy way to improve the performance of simulated annealing.

Acknowledgements:

This simulations were performed on the computers at the Institute for Molecular Science (IMS), Okazaki, Japan.

References

[1] S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, *Science*, **220**, 671 (1983).

[2] S. Geman and D. Geman, *IEEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-6** 721 (1984).

[3] H. Szu and R. Hartley, *Phys. Lett.* **A122** 157 (1987).

[4] T.J.P. Penna, *Phys. Rev. E* **51**, R1 (1995).

[5] I. Andricioaei and J.E.Straub, *Phys. Rev. E* **53** R3055 (1996).

[6] C. Tsallis, *J. Stat. Phys.* **52**, 479 (1988).

[7] D.A. Stariolo and C. Tsallis, *Annual Reviews of Computational Physics II*, edited by D. Stauffer (World Scientific, Singapore, 1995), p. 343.

[8] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, *J. Comp. Chem.* **4** 187 (1983).

[9] M.J. Sipple, G. Némethy, and H.A. Scheraga, *J. Phys. Chem.* **88**, 6231 (1984), and references therein.

[10] H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, and T. Kikuchi, *Chem. Lett.* **1991**, 213 (1991); Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, *Protein Engineering* **4**, 639 (1991).

[11] U.H.E. Hansmann and Y. Okamoto, *J. Phys. Soc. Jpn.* **63**, 3945 (1994); *Physica A* **212**, 415 (1994).

[12] Y. Okamoto, T. Kikuchi, and H. Kawai, *Chem. Lett.* **1992**, 1275 (1992).

[13] U.H.E. Hansmann and Y. Okamoto, "Numerical Comparisons of Three Recently Proposed Algorithms in the Protein Folding Problem", *J. Comp. Chem.*, in press.

Table 1:

Results for simulated annealing simulations using Tsallis weights as defined in Ref. [4]. For each value of q 20 independent runs of 50,000 sweeps were performed. The start temperature was $T_{ST} = 1000\ K$. $q = 1$ indicates results from regular simulated annealing which were taken from previous work in Ref. [11]. n_G is the number of runs where a ground state configuration was found. $< E_{Low} >$ is the average over 20 runs of the lowest energies obtained in each run. The standard deviation of this quantity is given in parentheses.

	$T_F = 50\ K$		$T_F = 1\ K$	
q	n_G	$< E_{Low} >$	n_G	$< E_{Low} >$
2.00	0/20	-6.0 (0.9)	2/20	-9.2 (1.3)
1.75	0/20	-7.5 (1.0)	1/20	-9.5 (1.2)
1.50	3/20	-9.6 (1.1)	1/20	-8.7 (1.3)
1.25	5/20	-9.5 (1.2)	4/20	-9.7 (1.4)
1.00	6/20	-10.0 (1.3)	8/20	-10.0 (2.2)
0.75	6/20	-10.2 (1.3)	7/20	-9.9 (1.9)
0.50	5/20	-10.2 (1.3)	6/20	-9.9 (1.6)
0.25	5/20	-10.2 (1.3)	8/20	-10.0 (1.7)
0.00	8/20	-10.4 (1.3)	2/20	-9.0 (1.8)
-0.50	5/20	-9.3 (1.8)	5/20	-9.1 (1.9)
-1.00	4/20	-9.5 (1.8)	5/20	-9.6 (1.9)
-2.00	1/20	-8.6 (1.8)	1/20	-8.1 (1.9)

Table 2:

Results for simulated annealing simulations using Tsallis weights as defined in Ref. [5]. For each annealing schedule characterized by the choice of start temperature T_{ST} and final temperature T_F 20 independent runs of 50,000 sweeps were performed. The results are compared with that of regular simulated annealing runs. n_G is the number of runs where a ground state configuration was found. $< E_{Low} >$ is the average over 20 runs of the lowest energies obtained in each run. The standard deviation of this quantity is given in parentheses.

T_{ST}/K	T_{FI}/K	Regular Simulated Annealing		Simulated Annealing Version of Ref. 5	
		n_G	$< E_{Low} >$	n_G	$< E_{Low} >$
1000	50	6	-10.0 (1.3)	0	-7.9 (1.8)
1000	1	8	-10.0 (2.2)	0	-7.9 (1.3)
500	50	8	-10.5 (1.3)	3	-8.3 (1.8)
500	1	2	-9.3 (1.3)	1	-8.1 (1.6)
300	50	1	-9.8(1.2)	1	-8.3 (1.4)
300	1	3	-9.6(1.4)	5	-8.9 (2.0)

Table 3:

Results for simulated annealing simulations using Tsallis weights as defined in Eq. 12. For each annealing schedule characterized by the choice of start temperature T_{ST} and final temperature T_F, 20 independent runs of 50,000 sweeps were performed. The results are compared with that of regular simulated annealing runs. n_G is the number of runs where a ground state configuration was found. $< E_{Low} >$ is the average over 20 runs of the lowest energies obtained in each run. The standard deviation of this quantity is given in parentheses.

T_{ST}/K	T_{FI}/K	Regular Simulated Annealing		New Simulated Annealing Version	
		n_G	$< E_{Low} >$	n_G	$< E_{Low} >$
1000	50	6	-10.0 (1.3)	7	-10.7 (0.9)
1000	1	8	-10.0 (2.2)	7	-10.7 (1.3)
500	50	8	-10.5 (1.3)	12	-11.1 (0.9))
500	1	2	-9.3 (1.3)	11	-10.9 (1.3))
300	50	5	-10.1 (1.3)	13	-11.0 (0.9)
300	1	3	-9.6 (1.4)	11	-11.0 (1.1)